

## Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*

(signal transduction/osmoregulation/phosphate regulation/*Agrobacterium* virulence/chemotaxis)

B. TRACY NIXON\*†‡, CLIVE W. RONSON\*§¶, AND FREDERICK M. AUSUBEL\*†

\*Department of Genetics, Harvard Medical School, and †Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114; and §Grasslands Division, Department of Scientific and Industrial Research, Private Bag, Palmerston North, New Zealand

Communicated by Philip Leder, July 10, 1986

**ABSTRACT** We report that the *ntrB* and *ntrC* proteins of *Bradyrhizobium* sp. [*Parasponia*] strain RP501 share homology with other regulatory proteins. There is extensive conservation of C-terminal regions between products of RP501 *ntrB*; *Klebsiella pneumoniae* *ntrB*; *Escherichia coli* *envZ*, *cpxA*, and *phoR*; *Agrobacterium tumefaciens* *virA*; and, to a lesser extent, *E. coli* *cheA*. There is also extensive conservation of N-terminal regions between products of RP501 *ntrC*; *K. pneumoniae* *ntrC*; *E. coli* *ompR*, *sfrA*, *phoB*, *cheY* and *cheB*; *Salmonella typhimurium* *cheB* and *cheY*; *Bacillus subtilis* *spo0A* and *spo0F*; and *A. tumefaciens* *virG*. We propose that these regulatory genes comprise two-component regulatory systems that evolved from a common ancestral system that involved transduction of information about the status of the environment by one protein domain (the C-terminal regions conserved among *ntrB*, *envZ*, etc.) to a second one (the N-terminal region conserved among *ntrC*, *ompR*, etc.). The *ntrC*-set protein then acts upon a specific responding mechanism, typically as a transcriptional activator but also as an effector of the maturation of outer membrane proteins or as a modulator of the direction of flagella rotation.

In the enteric bacteria *Escherichia coli*, *Salmonella typhimurium*, and *Klebsiella pneumoniae*, the products of three genes, *rpoN* (also called *ntrA* and *glnF*), *ntrB* (also called *glnL*), and *ntrC* (also called *glnG*) regulate the expression of several genes involved in the assimilation of nitrogen (reviewed in refs. 1 and 2). The *ntrC* protein acts either as a transcriptional activator or repressor depending on the promoter being regulated (refs. 1 and 3 and references cited therein), the *rpoN* protein is a  $\sigma$  factor (4, 5), and the *ntrB* protein modifies the functional state of *ntrC* product in response to the intracellular ratio of 2-ketoglutarate to glutamine, a signal of nitrogen availability (6–8). Information about the ratio is transmitted to the *ntrB* protein by the products of the *glnD* and *glnB* genes (6, 7).

In *K. pneumoniae*, transcription of the *nifA* gene is activated by *ntrC* and *rpoN* proteins under conditions of nitrogen limitation; subsequently, *nifA* and *rpoN* proteins activate transcription of the remaining *nif* genes (reviewed in ref. 9). A variety of nitrogen-fixing species of nonenteric bacteria, *Rhizobium meliloti* for example (10), also appear to have homologues of the *ntrC* and *nifA* genes, but homologues of *ntrB* have not been reported.

In this paper we report the cloning and sequence analysis of putative *ntrB* and *ntrC* genes of *Bradyrhizobium* sp. [*Parasponia*] (in this paper designated "*B. parasponiae*") strain RP501 isolated from nodules of *Parasponia rigida*, a nonlegume. We also report that the products of these *ntr*

genes share homology with those of a variety of two-component regulatory systems that control responses to environmental stimuli.

### MATERIALS AND METHODS

*B. parasponiae* RP501 (provided by J. Tjepkema) was used to isolate a variant resistant to streptomycin, RP501S, that was routinely grown at 32°C in TY medium supplemented with streptomycin (1.0 mg/ml).

Plasmids pGln53Y (11) and pRmC3.8R (10) provided *E. coli* and *R. meliloti* *ntrC* DNA, respectively, for hybridization experiments. Southern hybridizations were performed in 10% (wt/vol) dextran sulfate/50% (vol/vol) formamide at 42°C or 37°C. Nick-translations (12) and random-primer extensions (13) were carried out as described.

Libraries of RP501S DNA were constructed in cosmid vectors pHCT9 (14) and pOCA7.9 (Olszewski and F.M.A., unpublished results). Fragments of *B. parasponiae* DNA coding for putative *ntrB* and *ntrC* genes were recloned in pUC13 (15). Deletions of these DNA fragments generated by BAL-31 digestions (16), which together spanned the DNA of interest and provided both strands for analysis, were cloned in M13 vectors mp18 and mp19 (15) and sequenced by the Sanger dideoxy method as described (17). The 5' ends of open reading frames were defined by usage of preferred codons, presence of probable ribosome binding sites, and homology to *K. pneumoniae* *ntrB* and *ntrC* products. The NBRF program Pirsearch was used to scan the Dayhoff data base for protein homology, and the NBRF program Piralign (18) was used to calculate the relatedness of two sequences by determining the number of standard deviations separating the maximum alignment from the average of 100 randomized alignments. In these comparisons the mutation data matrix was used with no bias for exact homology, and a penalty of 16 was assigned to introduced gaps.

### RESULTS

**Cloning of *B. parasponiae* DNA Homologous to *ntrC*.** When  $^{32}$ P-labeled *E. coli* and *R. meliloti* *ntrC* probes from pGln53Y and pRmC3.8R, respectively, were hybridized to genomic *B. parasponiae* DNA using low stringency conditions, a homologous 1.7-kilobase EcoRI fragment was detected (data not shown). Since *E. coli* strains, including a presumptive *ntrC* deletion strain YMC11 (19), contained *ntrC* hybridizing sequences, *B. parasponiae* DNA was digested with EcoRI and size-fractionated on a sucrose gradient, and the fractions

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

\*Present address: Dept. of Molecular and Cell Biology, Pennsylvania State University, University Park, PA 16802.

†Present address: Biotechnica International, 85 Bolton Street, Cambridge, MA 02140.

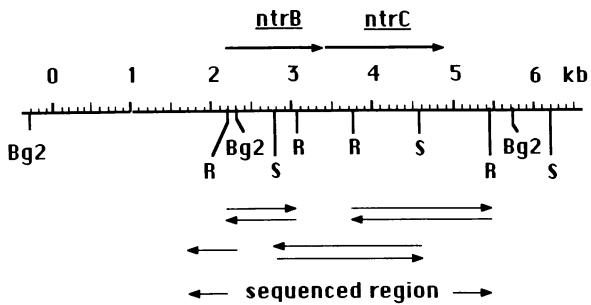


FIG. 1. Physical map and sequencing strategy of *B. parasponiae* DNA that hybridized to *E. coli* and *R. meliloti* *ntrC* genes. The indicated fragments were subcloned into pUC13 and sequenced (see Fig. 2), allowing the putative *ntrB* and *ntrC* genes to be positioned as shown (R, *EcoRI*, S, *Sal I*, Bg2, *Bgl II*).

most enriched for *ntrC*-homologous DNA were identified by Southern hybridization and cloned into pUC13. Hybridization of Southern blots to *ntrC* probe identified a recombinant plasmid in the pUC13 library that contained the 1.7-kilobase *Eco*RI fragment homologous to *ntrC*.

**DNA Sequence Analysis of Putative *B. parasponiae* *ntrB* and *ntrC* Genes.** The 1.7-kilobase EcoRI fragment from *B. parasponiae* was used to screen *B. parasponiae* DNA libraries constructed using partial EcoRI digests. Eight clones that

hybridized to the *ntrC* probe were identified, pRPNtr1-4 (pHC79-derived) and pRPNtr5-8 (pOCAT7.9-derived). The *ntrC*-homologous region in each of these cosmids was mapped using hybridization of the 1.7-kilobase fragment to various restriction enzyme digests of the cosmids. Appropriate *ntrC*-homologous restriction fragments (Fig. 1) were subcloned from the cosmids for determining the nucleotide sequence of the *ntrC* region. In enteric bacteria *ntrB* is located directly upstream from *ntrC*; thus, when the orientation of the putative *B. paraspioniae* *ntrC* gene was known, restriction fragments containing upstream sequences were subcloned to determine by sequence analysis whether *B. paraspioniae* also contained an *ntrB*-like gene.

The sequence data (Fig. 2) revealed two open reading frames separated by 55 base pairs. The first one of 1131 base pairs had extensive homology with the *ntrB* gene product of *K. pneumoniae* (20); comparison with the Piralign algorithm gave a score of 26.53 standard deviation (SD) units with the homology spread throughout the proteins but strongest in the N- and C-terminal portions. The second open reading frame of 1440 base pairs had extensive homology with *ntrC* protein of *K. pneumoniae* (17); the Piralign score was 75.43 SD units, and the homology was throughout the proteins (Fig. 2).

**The *B. paraspioniae* *ntrB* and *ntrC* Gene Products Share Regions of Homology with Other Regulatory Proteins.** A search of the Dayhoff data base for proteins homologous to

FIG. 2. Sequence analysis of the *B. paraspioniae* DNA homologous to *ntrB* and *ntrC*. Piralign was used to compare the deduced products with *K. pneumoniae* *ntrB* and *ntrC* products. Exact conservation is indicated by double underlining, functional conservation by single underlining. Five gaps were introduced to align the *ntrB* gene products, four in the *K. pneumoniae* protein of 4, 2, 4, and 2 amino acids at positions 83, 185, 270, and 285, respectively, and one in the *B. paraspioniae* product of 1 residue at position 175. The *K. pneumoniae* *ntrC* gene product contained an additional residue at position 333. Asterisks denote probable ribosome binding sites.

RpntrB	87	PRIGGDRQVLDHVAPLTERPGHIVVMLQERTIAD-----KMDRQLTHRSAARSVIALAAMLAHEIKNPLSGIRGAQALLEQD	
KpntrB	82	TLVIDGRSHILSSTAQRQLPVEILEMAPMDNR-----RLSQEQRQHAAQDVLGRGLAHEIKNPLGGRLGAAQQLLSKA	
EenvZ	179	IRIQNRPLEVDELAHQALQVGKGIIPP-PLREYGASeVRSVTRAFAHNMAAGVKQGLADDRTLIMAGVSHDLRTPLTRIRLATEMMSEQ	
EcpxA	183	AWSLAKPARKLNKADEVAGQNLQRQPELEAGQPEFLAAGASFQNMVTALERMMTSQORLSDISHELRTPLTRLQLGTALLRRR	
EcpbR	159	TRDFSRPLNVLNTGRHLEIRVMPYTHKQLLM-----VARDVTQMHQLEGARRNNFANVSHELRTPLTVLOGYLEMNNEQ	
AtvirA	429	PRPCQGEIQLELATACLCHYIDVRRKQTECD-----VLLARRLEHQRLAEGTAGGIAHEFNNILGSILGHAELAQNS	
AtvirA	424	VQPASSEVOLMELAAGCVSHYVIRCKQTQRD-----ILERRLKAERLEAVTLAGGIAHEFNNILGVILGYA-MAQN	
RpntrB	165	ASSEDR--LLTRLICDEADRIIVTLDVRMVEFGDDRPVARGPVNTHSVLDHVKR-----LAQSGFARNVRFIEDYDPSLP	
KpntrB	160	LDPDL--MEYTKVIIEQADRLRNLVDRLR-GPQHPGMHVTESIKHVAERVVK-----LVSMELPDNVKVLVDYDPSLP	
EenvZ	264	DG-----YLAESINKDIIECNAIEQFIDYLRTQGEMPMEMADLNAVLEGIVIA-----AESGYEREIETALYPGS	
EcpxA	270	SGESK-----ELERIETEAQRQLDSMINNDLLVMSRNQKNALVSETIKANQLWSEVLDNAAEAEQMGRSLTVNFPGPW	
EcpbR	233	PLEGAVREKALHTMREQTQRMGELVKQQLLTLSKIEAAPTHLNNEKVDVPMMLRVER-----EAQTLSQKKQTFTEIDNG_L	
AtvirA	505	VSRTSVTRYIDYIISSSGDRAMLIIDQILTLSRKQERMKIPFESVSELVTEIAP-----LLRMALPPNIELSFRDQMQS	
AtvirA	499	LHRRTYARHYIDRINAESNRARLIIDQILALSRRRERTARPFLSALVREIAP-----SLRVALPSEVEVDFNISQAQM	
RpntrB	238	PVLANQDQLIQVFLNLVKNAAEAVADLGTDALIQLTTAFRPGVRLSPVPGKKSRSVSLPLEFCVKDNGSGVPEDLLPNLFDPFVT	
KpntrB	226	ELPHDPDQIEQVLLNIVRNQALPGEPEGEITLRTRTAFQ-----LTLHGVRVRLAARIDVEDNGPGIPSHLQDTLFYPMVSG	
EenvZ	331	EVKMHPLSKRRAVANMVVNAARYGNW-----IKVSSGTEPNRAWFQVEDDGPJIAPEQRKHLFQPFVRGD	
EcpxA	345	PLYGNPNALESALENIVRNALRYSHTK-----IEVGFADVKDGITITVDDDGPGVSPEDREQIFRPFYRTD	
EcpbR	301	KVSGNEDQRLRSAISNLVYNAVNHTPEGTH-----ITVWRQWPVPHGAEFSVEDNGPGIAPEHIPRITERFYRV	
AtvirA	579	VIEGSPLELQQVLLINICKNASQAMTANG-QIDIIISQAFLP-VKKILAHGVMPPGDYWLSSISDNGGGIPEAVLPHIEPFFFST	
AtvirA	573	IVEGNNPLEIEQILMNLCNKNAAEACIGTG-RIEDSVYRSFW-KHKVLANGTIPAGDYLILLSFEDNGGGIGQAALPHIEPFFFRT	
EcheA	?	.....LAKAASQGLTVSEQMSDDEVAML-IFAPGFS	
RpntrB	323	---KQ-TGSGLGLALVAKIVGDHGGIIECESQPRKTT-FRVLDAVQRRQATRPKQPR*	
KpntrB	310	---RE-GGTGLGLSIARSLIDQHSGKIEFTSWPQHTE-FSVYLPK*	
EenvZ	397	S-ARTISGTGLGLAIWORIVDNHNGMELGTSERGG--LSIRAWLPVPTRAQGTTKEG*	
EcpxA	411	EARDRESGGTGLGLAIWETAIQQRHGWKAEDSPGLGLRLIWLPYKRS*	
ExphoR	369	KARSRQTTGGSGLGLAIWVKHVNHHESRLNIESVTGKGTRFSVPIERLIAKNSD*	
AtvirA	661	--RARN-GGTGLGLASVGHISAFAGYIDVSSVTVGHGTRFDIYLPPSSKEVNPDSFFGRN..122 aa*	
AtvirA	655	--RAQC-GGTGLGLSTVGHVSAAMGFVDVISTVGRGTRFDIYLPTSAKKPVNSESFFGPE..124 aa*	
EcheA	?	EQVTDV-SGRGVGMODVVKRNQKMGHHEVQSKNGTGTIRILLPLTLAILDGMSVRADE..132 aa*	

FIG. 3. Comparison of C-terminal portions of gene products homologous to *B. parasponiae ntrB*. After using Piralign to compare protein sequences (Table 1), the number and location of gaps were minimized by manual editing. This procedure reduced the alignment significance for any given pair of sequences but highlighted the regions strongly conserved between all proteins. Shaded residues are conserved, defined as 50% or more amino acids belonging to one of the groups (L, I, V, M, F, Y, W), (P, T, A, G, S), (Q, N, E, D), or (H, K, R). Abbreviations are as in Table 1.

the *ntrB* gene product using the NBRF program Pirsearch yielded only a single protein that encoded by the *E. coli envZ* gene (Fig. 3). Like *ntrB*, *envZ* appears to be part of a two-component regulatory system, working in conjunction with the *ompR* gene to control expression and/or maturation of outer membrane proteins (21–23). Hence, we compared the *E. coli ompR* (24, 25) and *B. parasponiae ntrC* products and found that they share a highly conserved N-terminal domain of about 115 amino acids (Fig. 4). Similar conservation between *E. coli ompR* and *K. pneumoniae ntrC* products was independently discovered by Drummond *et al.* (26).

These observations prompted us to examine a number of other putative regulatory genes that are part of two-gene systems and that are involved in responses to environmental stimuli. The C-terminal regions of *ntrB* were found to share homology with *E. coli cpxA* (ref. 27; R. Weber and P. Silverman, personal communication), *phoR* (H. Shinagawa, personal communication), *A. tumefaciens virA* (B. LeRoux and E. Nester, personal communication), and possibly *E. coli cheA* (28) (Fig. 3). Similarly, the N-terminal portion of the *ompR* product that is homologous to the *B. parasponiae ntrC* protein has been shown to be conserved among the products of *Bacillus subtilis spoOA* (29, 30) and *spoOF* (30); *E. coli phoB*

(31), *sfrA* (*dye*; 32), *cheB* and *cheY* (28); *S. typhimurium cheB* and *cheY* (33); and *A. tumefaciens virG* (34) (Fig. 4). Despite the wide taxonomic diversity of the species from which these gene products are derived, analysis with Piralign (Table 1) showed that the degree of relatedness of any pair within the *ntrB* and *ntrC* sets is highly significant except for some of those containing *cheA*. Only partial sequence is available for *cheA*, and the most N-terminal portion is not homologous with regions shared by other members of the *ntrB* set.

## DISCUSSION

We have cloned and sequenced a region of the *B. parasponiae* genome that contains open reading frames homologous to products of the *K. pneumoniae* nitrogen assimilation regulatory genes *ntrB* and *ntrC*. The 55-base-pair region that separates the *B. parasponiae ntrB* and *ntrC* genes is devoid of obvious transcription termination signals, suggesting that the genes are contiguous members of a nitrogen regulation operon.

Many of the regulatory proteins we found to be homologous to *ntrB* and *ntrC* products (Figs. 3 and 4) function in pairs to regulate gene expression in response to environment-

RpntrC	10	MPAGSILVADDDTAIRTVLNQALSRAGY	20	EVRLTG.NAATLWRWVSSQGEGLDVITDVVMPDE	30		40		50		60	
KpntrC		MQRGIAWIVDDDSIRRWLERALTGAGL		SCTTFS.GNEVLDALTTKTPDVLSSDIRMPGM								
EcoMpR		MQNQYKLNVYDDDMRLRALLERYLTQEGF		QVRVSA.NAEQMDRLLTRESFHLMVLDLMPGE								
Ecsfra		MQTPHILIVEDELVTRNLTKSIEFAEGY		DYFEAT.DGAEMHQILSEYDINLVINDINLPKG								
EcpbR		MARRILVYVEADIPEIRVMCFVLEQNGF		..EPLKHVLLYDDDVAMRHLLIEYLTIHAF.KVTAVA								
AtvirG		Q.PVVEAE.DYDSAQNQLNEPWPDLLLLDWMLPGG		DSTQFTRVLSSATVDVYVVDLNLVRE								
EcChey		..VALMFLTGDRNEVDTKILGLEIGADDYITTRPFSPRELVARLKV		AEDGVDALNKLQAGGYGFIISDWNMPNM								
EcCheb		SGIQFKHLKRESMTRDIPVVMLTARGEEEDRVRGLETGADDYITTRPFSPRELVARLKV		MNEKII1LIVDDQYGIIRELLNEVFNKEGY.QTFQAA.NGLQALDIDTVKERPDVLVLLDKMPGM								
BsSpoOF		MEKIKVSVADDMRELVSLLSEYIEGQEDMEVI		MEKIKVSVADDMRELVSLLSEYIEGQEDMEVI								
BsSpoOA		100	GVAYNGQECLSLFKEKDPDVVLVLDTIMPHE									
RpntrC	70	NAFDLLPKRIKKMRP	80	..NLPVIVMSAQNTFMTAIRPSERGAYLKPFPDLKELITIVGRA	90		100		110			
KpntrC		DGLALLKQIKQRHP..MLPVI		..IMTAHSDDAASVAYQQGAFDYLKPFPDLDEAVALVDR								
EcoMpR		DGLSICRRLRSQSN..PMP		..IIVMTAKGEEDRIVGLEIYGADDYIPKPFNPRLLAIRAV								
Ecsfra		NGLLLARELRLREQAN..		..VALMFLTGDRNEVDTKILGLEIGADDYITTRPFSPRELVARLKV								
EcpbR		SGIQFKHLKRESMTRDIPVVMLTARGEEEDRVRGLETGADDYITTRPFSPRELVARLKV		..DGELEI								
AtvirG		..DGELEI		..KQVLLAASKD..IPITIISGDRLEETDKVVALELGASDFIAKPKSIREFLARIRVA								
EcChey		..DGELEI		..KQVLLAASKD..IPITIISGDRLEETDKVVALELGASDFIAKPKSIREFLARIRVA								
EcCheb		..DGLDFLEKLMLRPPM		..VYVMSLSTGKGSEVTL.RALELGAI								
BsSpoOF		..VYVMSLSTGKGSEVTL.RALELGAI		..DGELEI								
BsSpoOA		..DGELEI		..KQVLLAASKD..IPITIISGDRLEETDKVVALELGASDFIAKPKSIREFLARIRVA								

FIG. 4. Comparison of N-terminal portions of gene products homologous to *B. parasponiae ntrC* (since *cheB* and *cheY* from *S. typhimurium* and *E. coli* are strongly homologous, only the *E. coli* sequences are shown). Alignments were prepared as in Fig. 3, except conservation was assumed when 40% or more of the sequences contained similar amino acids. Abbreviations are as in Table 1.

Table 1. Pair-wise comparisons of regulatory proteins homologous to *ntrB* and *ntrC*

	<i>RpntrB</i>	<i>KpntrB</i>	<i>AtvirA*</i>	<i>AtvirA†</i>	<i>EcenvZ</i>	<i>EcphoR</i>	<i>EccpxA</i>	<i>EccheA</i>	
	97.29	27.99	14.53	9.48	10.48	9.26	6.10	2.24	<i>RpntrB</i>
		116.81	15.75	8.79	10.38	9.61	9.81	2.54	<i>KpntrB</i>
<i>RpntrC</i>	62.89			99.04	67.66	6.53	9.85	11.65	6.61
<i>KpntrC</i>	23.22	61.15			108.36	6.62	9.80	13.92	5.96
<i>EcompR</i>	27.62	13.38	70.95			110.86	17.84	20.59	8.62
<i>EcsfrA</i>	13.65	12.52	28.11	58.40			111.63	18.59	5.58
<i>EcphoB</i>	12.45	9.26	25.31	23.77	58.95			95.72	5.09
<i>AtvirG</i>	10.42	13.30	28.75	20.41	20.05	57.35			46.50
<i>EccheY</i>	8.75	14.71	15.37	10.21	17.89	8.23	55.10		
<i>EccheB</i>	6.93	7.00	6.26	9.56	7.89	8.19	5.18	53.80	
<i>BsspooF</i>	18.22	18.96	14.21	14.34	18.23	8.74	8.64	7.34	51.08
<i>BsspooA</i>	6.16	11.42	13.38	7.90	13.71	7.16	7.84	14.82	11.31
<i>RpntrC</i>		<i>KpntrC</i>	<i>EcompR</i>	<i>EcsfrA</i>	<i>EcphoB</i>	<i>AtvirG</i>	<i>EccheY</i>	<i>EccheB</i>	<i>BsspooF</i>
									<i>BsspooA</i>

Piralign was used to determine the relatedness (SD units) of regulatory proteins to the conserved C-terminal regions of *B. parasponiae ntrB* (top, see Fig. 3) and conserved N-terminal domain of *B. parasponiae ntrC* (bottom, see Fig. 4). Rp, RP501; Kp, *K. pneumoniae*; Ec, *E. coli*; At, *A. tumefaciens*; Bs, *B. subtilis*; \*, Wide or, †, limited host range.

tal stimuli such as nutrient limitation (*ntrB/ntrC*, *phoR/phoB*), altered osmolarity (*envZ/ompR*), and plant exudate (*virA/virG*). Except for *ntrB*, which is the only member of its set clearly responsive to an intracellular signal, the products of all members of the *ntrB* set have hydrophobic profiles (data not shown) consistent with transmembrane structure in the N-terminal portion and cytoplasmic location of the C-terminal domain, analogous to the chemosensory transducing proteins encoded by *tar*, *tsr*, and *trg* (35–37). Furthermore, as shown for the *ntrB/ntrC* products (6), it has been suggested that *envZ*, *phoR*, and *virA* products modify their respective partners, the *ompR*, *phoB*, and *virG* products (22, 23, 32, 34).

On the basis of the above observations, we propose a model for transduction of environmental signals by these pairs of regulatory gene products. As suggested for chemo-receptors (35–37), the nonconserved N-terminal domains of the *ntrB* set of proteins perceive environmental signals and transmit them to the conserved cytoplasmic, C-terminal portions through allosteric modifications. The C-terminal portion of the *ntrB* set protein then interacts with or modifies the N-terminal portion of its corresponding *ntrC*-set protein, with the modification affecting activation or repression through an allosteric effect(s) on the C-terminal domain(s) of the *ntrC*-set protein or other proteins involved in the response. In support of this model, evidence has demonstrated that *ntrB* product is a kinase/phosphatase that phosphorylates the *ntrC* protein when ammonia is limiting and dephosphorylates it in response to increased ammonia availability (6). This finding provides a physical mechanism for the information transfer between *ntrB* and *ntrC* products and raises the possibility that other members of the *ntrB* set may modulate the function of their counterpart protein in the *ntrC* set in a similar manner.

If this model is correct, it should be feasible to isolate mutant proteins with regulatory domains locked in active conformation. Such mutants may have been described for *ntrB* (38, 39), *ntrC* (1), *cpxA* (27, 40), and *ompR* (23, 24). Obtaining pseudorevertants of these and similar mutants by alterations in counterpart domains would demonstrate interaction between the proposed regulatory regions. Moreover, given the strong conservation between the various members of the *ntrB* and *ntrC* sets in the domains that we propose are responsible for information transfer, it is possible that one member of the *ntrB* set may modify the function of more than one member of the *ntrC* set. Indeed *cpxA* mutations can affect the expression of *ompF* (41), and some mutants in *envZ* exhibit pleiotropic effects upon expression of several genes, including *phoA* (23). Such interactions may account for the ability of *ntrB* mutants of *E. coli* to activate *ntrC* slowly in

response to nitrogen limitation (8). In similar fashion, our model also predicts that it may be possible to bring expression of *nif* genes under control of *phoR-phoB* regulation, for example, by replacing the *phoB* C-terminal domain with that of *ntrC*.

Among the transcriptional activators, *ompR*, *sfrA*, *phoB*, and *virG* products share homology over their entire length, while *ntrC* and *spo0A* products share only the N-terminal domain. Outside of the conserved N-terminal domain, the remainder of *ntrC* product is highly conserved with the analogous regions of *nifA* proteins (17, 26). The *ntrC* and *nifA* products share several common features in their mode of action, including requirements for an upstream binding site in regulated promoters (3, 42, 43) and the  $\sigma$  factor encoded by *rpoN* (11, 44) for transcriptional activation. Hence the common portions of *ntrC* and *nifA* proteins are probably concerned with these features, and indeed both proteins contain a consensus DNA-binding motif within their last 25 amino acids (26). It is possible that the C-terminal portions of *ompR*, *phoB*, *virG*, and *sfrA* proteins may also be involved in interacting with a  $\sigma$  factor or the core of RNA polymerase since, at least for the first three, the promoters they regulate do not contain consensus –35 regions (23, 31, 45).

The lack of conservation between the N-terminal domains of the *K. pneumoniae ntrC* and *nifA* products may account for their different sensitivities to environmental stimuli. Whereas *ntrB* product modulates the activity of *ntrC* product in response to nitrogen status, *K. pneumoniae nifL* product inhibits *nifA* product activity in response to oxygen tension or low amounts of combined nitrogen (46). It seems likely that in this species the N-terminal part of the *nifA* protein interacts with *nifL* protein and that the *nifL* protein will not contain the *ntrB*-set conserved sequence. Furthermore, although there is similarity between the N-terminal portions of *K. pneumoniae* and *R. meliloti nifA* gene products (alignment score of 3 SD units), it is possible that the low degree of conservation reflects evolution of the N-terminal region of the rhizobial *nifA* product to respond to some aspect of the symbiotic environment.

Products of the chemotaxis genes *cheA*, *cheB*, and *cheY* are also involved in response to environmental signals, although the response is not thought to involve activation or suppression of transcription. The *cheB* protein is a methyl esterase that is activated in response to repellent stimuli (47, 48) and acts during adaptation to demethylate the transducing proteins (49, 50). Interestingly, proteolytic cleavage or genetic deletion of the N-terminal domain of *cheB* product results in a 15-fold increase in specific activity of the enzyme (51), suggesting that the conserved N-terminal domain can indeed modulate the activity of the remainder of the protein

as predicted in the above model. The cheY product is thought to interact directly with the basal structure of the flagella to modulate swimming behavior in response to unknown signals from the receptors (52, 53). The homology of cheB and cheY products to the ntrC-set conserved domain suggests that the functions of these proteins may be modulated by an ntrB-set counterpart, possibly cheA product that has significant homology to most members of the ntrB set. It may be important that the most N-terminal of available sequence for the cheA product is not homologous to regions shared by the other proteins.

Finally, the two-component regulatory model proposed above predicts that the unpaired genes in the ntrC set, such as the *spo0* genes that respond to nutrient deprivation, will have partners in the ntrB set. Since *sfrA* and *cpxA* both regulate expression of *tra* genes (32, 54), they may form a pair responsive to the presence of F<sup>-</sup> cells. The model also predicts that other two-component regulatory systems that respond to environmental stimuli, such as *xylS/xylR* of *Pseudomonas putida* (ref. 55 and references cited therein), may have homology to the domains reported here.

We thank D. Comeau for helpful discussion and B. LeRoux, S. Winans, E. Nester, R. Weber, P. Silverman, H. Shinagawa, and B. Magasanik for providing data prior to publication. This work was supported by a grant from Hoechst AG.

- Kustu, S., Sei, K. & Keener, J. (1986) in *Regulation of Gene Expression—25 Years On*: Society for General Microbiology, eds. Booth, I. & Higgins, C. F. (Cambridge Univ. Press, Cambridge, U.K.), pp. 139–154.
- Magasanik, B. (1982) *Annu. Rev. Genet.* **16**, 135–168.
- Reitzer, L. J. & Magasanik, B. (1986) *Cell* **45**, 785–792.
- Hirschman, J., Wong, P.-K., Sei, K., Keener, J. & Kustu, S. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7525–7529.
- Hunt, T. P. & Magasanik, B. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8453–8457.
- Ninfa, A. J. & Magasanik, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5909–5913.
- Bueno, R., Pahel, G. & Magasanik, B. (1985) *J. Bacteriol.* **164**, 816–822.
- Reitzer, L. J. & Magasanik, B. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1979–1983.
- Ausubel, F. M. (1984) in *Microbial Development* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 275–298.
- Ausubel, F. M., Buikema, W. J., Earl, C. D., Klingensmith, J. A., Nixon, B. T. & Szeto, W. W. (1985) in *Nitrogen Fixation Research Progress*, eds. Evans, H. J., Bottomley, P. J. & Newton, W. E. (Nijhoff, The Hague), pp. 165–171.
- Ow, D. W. & Ausubel, F. M. (1983) *Nature (London)* **301**, 307–313.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 110–112.
- Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
- Hohn, B. & Collins, J. (1980) *Gene* **11**, 291–298.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103–119.
- Poncz, M., Solowiejczyk, B., Ballantine, M., Schwartz, E. & Surrey, S. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4298–4302.
- Buikema, W. J., Szeto, W. W., Lemley, P. V., Orme-Johnson, W. H. & Ausubel, F. M. (1985) *Nucleic Acids Res.* **13**, 4539–4555.
- Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983) *Methods Enzymol.* **91**, 524–545.
- Backman, K., Chen, Y.-M. & Magasanik, B. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3743–3747.
- MacFarlane, S. A. & Merrick, M. (1985) *Nucleic Acids Res.* **13**, 7591–7606.
- Hall, M. N. & Silhavy, T. J. (1979) *J. Bacteriol.* **140**, 342–350.
- Hall, M. N. & Silhavy, T. J. (1981) *J. Mol. Biol.* **146**, 23–43.
- Ramakrishnan, G., Comeau, D. E., Ikenaka, K. & Inouye, M. (1986) in *Bacterial Outer Membranes as Model Systems*, ed. Inouye, M. (Wiley, New York), in press.
- Nara, F., Matsuyama, S., Mizuno, T. & Mizushima, S. (1986) *Mol. Gen. Genet.* **202**, 194–199.
- Comeau, D. E., Ikenaka, K., Tsung, K. & Inouye, M. (1985) *J. Bacteriol.* **164**, 578–584.
- Drummond, M., Whitty, P. & Wootton, J. (1986) *EMBO J.* **5**, 441–447.
- Albin, R., Weber, R. & Silverman, P. M. (1986) *J. Biol. Chem.* **261**, 4698–4705.
- Mutoh, N. & Simon, M. I. (1986) *J. Bacteriol.* **165**, 161–166.
- Ferrari, F. A., Trach, K., LeCoq, D., Spence, J., Ferrari, E. & Hoch, J. A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2647–2651.
- Trach, K. A., Chapman, J. W., Piggot, P. J. & Hoch, J. A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7260–7264.
- Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1986) *J. Mol. Biol.* **190**, 37–44.
- Drury, L. S. & Buxton, R. S. (1985) *J. Biol. Chem.* **260**, 4236–4242.
- Stock, A., Koshland, D. E. & Stock, J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7989–7993.
- Winans, S. C., Ebert, P. R., Stachel, S. E., Gordon, M. P. & Nester, E. W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, in press.
- Krikos, A., Mutoh, N., Boyd, A. & Simon, M. I. (1983) *Cell* **33**, 615–622.
- Bollinger, J., Park, C., Harayama, S. & Hazelbauer, G. L. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3287–3291.
- Boyd, A., Kendall, K. & Simon, M. I. (1983) *Nature (London)* **301**, 623–626.
- Chen, Y.-M., Backman, K. & Magasanik, B. (1982) *J. Bacteriol.* **150**, 214–220.
- Pahel, G., Zelenetz, A. D. & Tyler, B. M. (1978) *J. Bacteriol.* **133**, 139–148.
- Albin, R. & Silverman, P. M. (1984) *Mol. Gen. Genet.* **197**, 272–279.
- McEwen, J., Sambucetti, L. & Silverman, P. M. (1983) *J. Bacteriol.* **154**, 375–382.
- Buck, M., Miller, S., Drummond, M. & Dixon, R. (1986) *Nature (London)* **320**, 374–378.
- Better, M., Ditta, G. & Helinski, D. R. (1985) *EMBO J.* **4**, 2419–2424.
- Merrick, M. J. (1983) *EMBO J.* **2**, 39–44.
- Das, A., Stachel, S., Ebert, P., Allenza, P., Montoya, A. & Nester, E. (1986) *Nucleic Acids Res.* **14**, 1355–1364.
- Buchanan-Wollaston, V. & Cannon, F. (1984) in *Advances in Nitrogen Fixation Research*, eds. Veeger, C. & Newton, W. E. (Nijhoff/Junk, The Hague), p. 732.
- Kehry, M. R., Doak, T. G. & Dahlquist, F. W. (1984) *J. Biol. Chem.* **259**, 11828–11835.
- Springer, M. S., Goy, M. F. & Adler, J. (1979) *Nature (London)* **280**, 279–284.
- Stock, J. B. & Koshland, D. E., Jr. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3659–3663.
- Hayashi, H., Koiwai, O. & Kozuka, M. (1979) *J. Biochem.* **85**, 1213–1223.
- Simms, S. A., Keane, M. G. & Stock, J. (1985) *J. Biol. Chem.* **260**, 10161–10168.
- Parkinson, J. S., Parker, S. R., Talbert, P. B. & Houts, S. E. (1983) *J. Bacteriol.* **155**, 265–274.
- Clegg, D. O. & Koshland, D. E., Jr. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5056–5060.
- Cuozzo, M. & Silverman, P. M. (1986) *J. Biol. Chem.* **261**, 5175–5179.
- Inouye, S., Nakazawa, A. & Nakazawa, T. (1985) *J. Bacteriol.* **163**, 863–869.